

## PAGE ANALYSIS SYSTEM

Publication number: JP8235349

Publication date: 1996-09-13

Inventor: SHIN YAN WANGU

Applicant: CANON KK

Classification:

- international: G06K9/20; G06T1/00; G06T11/60; H04N1/40;  
H04N1/41; G06K9/20; G06T1/00; G06T11/60;  
H04N1/40; H04N1/41; (IPC1-7): G06T1/00; G06K9/20;  
G06T11/60

- European: G06K9/20L; H04N1/40L; H04N1/41

Application number: JP19950293123 19951110

Priority number(s): US19940338781 19941110

Also published as:

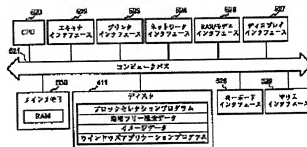
EP0712088 (A2)  
US6014458 (A1)  
US5987171 (A1)  
EP0712088 (A3)  
EP0712088 (B1)

more &gt;&gt;

Report a data error here

## Abstract of JP8235349

**PURPOSE:** To execute high-precision block selection by monitoring the data volume and the distortion of a document as the processing object of block selection processing and indicating execution/non-execution of the processing based on the monitor result. **CONSTITUTION:** A main memory 530 is connected to a computer bus 521 and is provided with a RAM. This RAM is the work area to execute various processings (for example, the block selection processing) executed by a CPU 520. With respect to a deficiency of the storage capacity in a first storage area, a first error code indicating this deficiency is outputted when picture data exceeding the one-page portion of a document consisting of plural pages is read in. With respect to a deficiency of the storage capacity in a second storage area, a second error code indicating this deficiency is outputted when the second storage area has no space to store the effective processing result of the block selection processing which can be read and stored in the second storage area.



Data supplied from the esp@cenet database - Worldwide

**Family list****11** family members for: **JP8235349**

Derived from 7 applications

Back to JP823

- 1 System for designating document direction**  
**Inventor:** WANG SHIN-YWAN (US) **Applicant:** CANON KK (JP)  
**EC:** G06K9/20L; H04N1/40L; (+1) **IPC:** G06K9/20; G06T1/00; G06T11/60 (+8)  
**Publication info:** **DE69532847D D1** - 2004-05-13
- 2 System for designating document direction**  
**Inventor:** WANG SHIN-YWAN (US) **Applicant:** CANON KK (JP)  
**EC:** G06K9/20L; H04N1/40L; (+1) **IPC:** G06K9/20; G06T1/00; G06T11/60 (+8)  
**Publication info:** **DE69532847T T2** - 2005-03-17
- 3 Page analysis system**  
**Inventor:** WANG SHIN-YWAN (US) **Applicant:** CANON KK (JP)  
**EC:** G06K9/20L; H04N1/40L; (+1) **IPC:** G06K9/20; G06T1/00; G06T11/60 (+8)  
**Publication info:** **EP0712088 A2** - 1996-05-15  
**EP0712088 A3** - 1998-01-07  
**EP0712088 B1** - 2004-04-07
- 4 Document analysis system and method**  
**Inventor:** WANG SHIN-YWAN (US) **Applicant:** CANON KK (JP)  
**EC:** G06K9/20L2 **IPC:** G06K9/20; G06K9/20; (IPC1-7): G06K9/20  
**Publication info:** **EP1296278 A2** - 2003-03-26  
**EP1296278 A3** - 2003-04-02
- 5 PAGE ANALYSIS SYSTEM**  
**Inventor:** SHIN YAN WANGU **Applicant:** CANON KK  
**EC:** G06K9/20L; H04N1/40L; (+1) **IPC:** G06K9/20; G06T1/00; G06T11/60 (+10)  
**Publication info:** **JP3825820B2 B2** - 2006-09-27  
**JP8235349 A** - 1996-09-13
- 6 Page analysis system**  
**Inventor:** WANG SHIN-YWAN (US) **Applicant:** CANON KK (JP)  
**EC:** G06K9/20L; H04N1/40L; (+1) **IPC:** G06K9/20; G06T1/00; G06T11/60 (+8)  
**Publication info:** **US5987171 A** - 1999-11-16
- 7 System for designating document direction**  
**Inventor:** WANG SHIN-YWAN (US) **Applicant:** CANON KK (JP)  
**EC:** G06K9/20L; H04N1/40L; (+1) **IPC:** G06K9/20; G06T1/00; G06T11/60 (+8)  
**Publication info:** **US6014458 A** - 2000-01-11

Data supplied from the esp@cenet database - Worldwide

(19)日本国特許庁 (J P)

## (12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平8-235349

(43)公開日 平成8年(1996)9月13日

(51)Int.Cl. <sup>8</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 T 1/00			G 0 6 F 15/62	3 3 0 D
11/60			G 0 6 K 9/20	3 4 0 L
G 0 6 K 9/20	3 4 0		G 0 6 F 15/62	3 2 5 A

審査請求 未請求 請求項の数18 O L (全 28 頁)

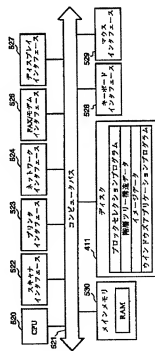
(21)出願番号	特願平7-293123	(71)出願人	000001007 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
(22)出願日	平成7年(1995)11月10日	(72)発明者	シン・ヤン ワング アメリカ合衆国 カリフォルニア州 92680, タスティン, マックチャールス ドライブ 2221
(31)優先権主張番号	0 8 / 3 3 8 7 8 1	(74)代理人	弁理士 大塚 康徳 (外1名)
(32)優先日	1994年11月10日		
(33)優先権主張国	米国 (U S)		

(54)【発明の名称】 ページ解析システム

## (57)【要約】

【課題】 ブロックセレクション処理の処理対象である原稿のデータ量と歪みを監視することで、処理精度が向上するブロックセレクション処理を実行できるページ解析システムを提供する。

【解決手段】 複数ページの原稿の1ページ分の画像データを解析するブロックセレクション処理を生成するページ解析システムにおいて、システムは、予め割り当てられた記憶領域に、その記憶容量を越える共通作業領域あるいは階層ツリー構造格納領域としてのデータが格納された場合にエラーコードを出力する。また、原稿の傾き角度を算出し、その算出された傾き角度が所定最大傾き角度を越える場合にエラーコードを出力する。



1

## 【特許請求の範囲】

【請求項1】 複数ページの画像データを解析するためのページ解析システムにおいて、該システムは、ブロックセレクション処理プログラムに従う処理を実行するために、複数ページの原稿の1ページ分の画像データを読み込み格納する第1記憶領域と、第1記憶領域に格納された画像データに対して実行されたブロックセレクション処理の処理結果を格納する第2記憶領域とを備え、該システムにおける記憶容量不足の指示の出力を行う出力方法であって、

前記第1記憶領域に、1ページ分の入力画像データを格納する記憶領域と前記入力画像データのブロックセレクション処理を実行するための記憶領域を割り当てる第1割当工程と、

前記第2記憶領域に、ブロックセレクション処理の複数の処理結果を格納するための記憶領域を割り当てる第2割当工程と、

前記複数ページの原稿の解析されるべき1ページ分の画像データが、前記第1記憶領域に割り当てられた記憶領域の記憶容量を越える場合に第1エラーコードを出力し、前記第2記憶領域に格納されたブロックセレクション処理の処理結果を利用するための記憶領域を増やすことができない場合に第2エラーコードを出力する出力工程とを備えることを特徴とする出力方法。

【請求項2】 前記ブロックセレクション処理の処理結果が格納される前記第2記憶領域は、階層ツリー構造を含むことを特徴とする請求項1に記載の出力方法。

【請求項3】 原稿1ページ分の画像データを解析するページ解析システムにおいて、前記原稿の少なくとも1つの領域を選択する選択方法であって、

原稿1ページ分の画像データを入力する入力工程と、前記画像データを用いて前記原稿ページを表示する表示工程と、

座標指定手段を用いて、解析すべき前記原稿ページの少なくとも1つの領域の座標を指定する指定工程と、前記指定工程で指定された前記原稿ページの少なくとも1つの領域内の画像データを抽出する抽出工程と、前記少なくとも1つの領域内の画像データを解析するために前記少なくとも1つの領域に対してブロックセレクション処理を実行する実行工程とを備えることを特徴とする選択方法。

【請求項4】 原稿1ページ分の画像データを解析するページ解析システムにおいて、前記原稿を解析するために少なくとも1つの方向を指定する指定方法であって、原稿1ページ分の画像データを入力する入力工程と、前記画像データの方向情報を獲得する獲得工程と、前記方向情報が獲得されたか否かを判定する判定工程と、

前記原稿を解析する解析工程を備え、

前記方向情報が獲得された場合、前記解析工程は該方向

2

情報によって指定される方向を用いて解析し、

前記方向情報が獲得されない場合、前記解析工程は予め設定されている所定方向を用いて解析することを特徴とする指定方法。

【請求項5】 前記所定方向は、水平方向と垂直方向が混合された方向であることを特徴とする請求項4に記載の指定方法。

【請求項6】 原稿1ページ分の画像データを解析し、ブロックセレクション処理によって解析された1つのルートノードと該ルートノードの子孫ノードを含む階層ツリー構造を生成するページ解析システムにおいて、前記原稿ページの傾き角度を獲得する獲得方法であって、前記原稿ページの画像データを入力する入力工程と、前記画像データに対し、前記画像データの傾き角度の算出を含むブロックセレクション処理を開始する開始工程と、

前記算出された傾き角度と所定最大傾き角度を比較する比較工程と、

前記算出された傾き角度が前記所定最大傾き角度よりも大きいか否かを判定する判定工程と、

前記傾き角度が前記所定最大傾き角度を越えない場合、前記画像データのブロックセレクション処理を完了する完了工程と、

前記傾き角度が前記所定最大傾き角度を越える場合、ブロックセレクション処理を中断する中断工程とを備えることを特徴とする獲得方法。

【請求項7】 前記傾き角度は、前記階層ツリー構造のルートノードに格納されることを特徴とする請求項6に記載の獲得方法。

【請求項8】 複数ページの画像データを解析するためのページ解析システムにおいて、該システムは、ブロックセレクション処理プログラムに従う処理を実行するために、複数ページの原稿の1ページ分の画像データを読み込み格納する第1記憶領域と、第1記憶領域に格納された画像データに対して実行されたブロックセレクション処理の処理結果を格納する第2記憶領域を備え、該システムにおける記憶容量不足の指示の出力を行う出力装置であって、

前記第1記憶領域に、1ページ分の入力画像データを格納する記憶領域と前記入力画像データのブロックセレクション処理を実行するための記憶領域を割り当て、前記第2記憶領域において、ブロックセレクション処理の複数の処理結果を格納するための記憶領域を割り当てる割当手段と、

前記複数ページの原稿の解析されるべき1ページ分の画像データが、前記第1割当手段に割り当てられた記憶領域の記憶容量を越える場合に第1エラーコードを出力し、前記第2記憶領域に格納されたブロックセレクション処理の処理結果を利用するための記憶領域を増やすことができない場合に第2エラーコードを出力する出力手

3

段とを備えることを特徴とする出力装置。

【請求項9】 前記ブロックセレクション処理の結果が格納される前記第2記憶領域は、階層ツリー構造を含むことを特徴とする請求項8に記載の出力装置。

【請求項10】 原稿1ページ分の画像データを解析するページ解析システムにおいて、前記原稿ページの少なくとも1つの領域を選択する選択装置であって、原稿1ページ分の画像データを入力する入力手段と、前記画像データを用いて前記原稿ページを表示する表示手段と、

解析すべき前記原稿ページの少なくとも1つの領域の座標を指定する座標指定手段と、

前記座標指定手段で指定された前記原稿ページの少なくとも1つの領域内の画像データを抽出する画像データ抽出手段と、

前記領域内の画像データを解析するために前記少なくとも1つの領域に対してブロックセレクション処理を実行する実行手段とを備えることを特徴とする選択装置。

【請求項11】 原稿1ページ分の画像データを解析するページ解析システムにおいて、前記原稿ページを解析するために少なくとも1つの方向を指定する指定装置であって、

原稿1ページ分の画像データを入力する入力手段と、前記画像データの方向情報を獲得する獲得手段と、前記方向情報が獲得されたか否かを判定する判定手段と、

前記原稿ページを解析する解析手段とを備え、前記方向情報が獲得された場合、前記解析手段は該方向情報によって指定された方向を用いて解析し、

前記方向情報が獲得されない場合、前記解析手段は予め設定されている所定方向を用いて解析することを特徴とする指定装置。

【請求項12】 前記所定方向は、水平方向と垂直方向が混合された方向であることを特徴とする請求項11に記載の指定装置。

【請求項13】 原稿1ページ分の画像データを解析し、ブロックセレクション処理によって前記原稿ページの1つのルートノードと該ルートノードの少なくとも1つの子孫ノードを含む階層ツリー構造を生成するページ解析システムにおいて、前記原稿ページの傾き角度を獲得する獲得装置であって、

前記原稿ページの画像データを入力する入力手段と、算出された傾き角度と所定最大傾き角度を比較する比較手段と、

前記算出された傾き角度が前記所定最大傾き角度より大きいと否かを判定する判定手段と、

入力された原稿ページの前記画像データに対して、傾き角度の算出を含むブロックセレクション処理を実行するブロックセレクション処理手段とを備え、

前記ブロックセレクション処理手段は、ブロックセ

4

クション処理を開始し、算出された傾き角度と所定最大傾き角度を比較するように前記比較手段を用い、前記算出された傾き角度が前記所定最大傾き角度を越えない場合、ブロックセレクション処理を完了させることを特徴とする獲得装置。

【請求項14】 前記傾き角度は、前記階層ツリー構造のルートノードに格納されることを特徴とする請求項13に記載の獲得装置。

【請求項15】 複数ページの原稿の画像データを解析するためのページ解析システムにおいて、ブロックセレクション処理プログラムによって解析されたページの階層ツリー構造が生成され、前記階層ツリー構造は1つのルートノードと該ルートノードの少なくとも1つの子孫ノードを含み、前記複数ページの原稿1ページの画像データの合成領域を指示する指示方法であって、前記複数ページの1ページ分の画像データを入力する入力手段と、

前記画像データを識別し、同じタイプの画像データを有するブロックを認識へ分けるようにブロックセレクション処理を準備する準備工程と、

同じタイプの画像データのブロックを他を持つ少なくとも1つのブロックを認識する認識工程と、

前記同じタイプの画像データを持つ複数のブロックを1つの合成ブロックに結合する結合工程と、

前記合成ブロックを、前記複数のブロックに対応する子孫ノードを持つルートノードとして前記階層ツリー構造に定時する定時工程とを備えることを特徴とする指示方法。

【請求項16】 複数ページの原稿の画像データを解析するためのページ解析システムにおいて、前記原稿ページを解析するための階層ツリー構造と、前記階層ツリー構造は1つのルートノードと該ルートノードの少なくとも1つの子孫ノードを含み、前記複数ページの原稿1ページの画像データの合成領域を指示する指示装置であって、

前記複数ページの1ページ分の画像データを入力する入力手段と、

前記画像データを識別し、同じタイプの画像データを有するブロックへ分けるブロックセレクション処理手段と、

同じタイプの画像データのブロックを他を持つ少なくとも1つのブロックを認識する認識手段と、

前記同じタイプの画像データを持つ複数のブロックを1つの合成ブロックに結合する結合手段と、

前記合成ブロックを、前記複数のブロックに対応する子孫ノードを持つルートノードとして前記階層ツリー構造を変更する変更手段とを備えることを特徴とする指示方法。

【請求項17】 1ページ分の原稿の画像データのブロックテンプレートに従って階層ツリー構造を生成するべ

5

ージ解析システムにおいて、前記階層ツリー構造は複数のノードを持ち、各ノードは原稿画像のブロックテンプレートにおける表されるブロック毎の画像データに対応しており、また、各ノードには前記原稿画像のブロックテンプレートの特徴を定義する特徴データが含まれ、該システムにおいて図形タイプの指示及び認識方法であって、

複数ページの原稿の1ページ分の画像データを入力する入力工程と、

前記画像データを識別し同じ画像タイプのブロック毎に分割するブロックセレクション処理を実行する実行工程と、

図形情報を含む画像データのブロックを認識する認識工程と、

前記ブロックに含まれる図形情報のタイプを判定する判定工程と、

前記判定工程は、図形情報がハーフトーン、線画、折れ線、未知のいずれかを判定し、

前記判定工程で判定される図形のタイプを、前記ブロックに対応する前記階層ツリー構造のノードに格納する格納工程とを備えることを特徴とする指示及び認識方法。

【請求項18】 1ページ分の原稿の画像データのブロックテンプレートに従って階層ツリー構造を生成するページ解析システムにおいて、前記階層ツリー構造は複数のノードを持ち、各ノードは原稿画像のブロックテンプレートにおける表されるブロック毎の画像データに対応しており、また、各ノードには前記原稿画像のブロックテンプレートの特徴を定義する特徴データが含まれ、該システムにおいて図形タイプの指示及び認識装置であって、

複数ページの原稿の1ページ分の画像データを入力する入力手段と、

前記画像データを識別し同じ画像タイプのブロック毎に分割し、分割されたブロックを認識するブロックセレクション処理と、

図形情報を含む画像データのブロックを認識する認識手段と、

前記ブロックに含まれる図形情報のタイプを判定する判定手段と、

前記判定手段は、図形情報がハーフトーン、線画、折れ線、未知のいずれかを判定し、

前記判定手段で判定される図形のタイプを、前記ブロックに対応する前記階層ツリー構造のノードに格納する格納手段とを備えることを特徴とする指示及び認識装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、ブロックセレクション処理を利用して、ページ単位で原稿を解析するページ解析システムに関するものである。特に、ブロックセレクション処理で表現される複数のパラメータ、例え

6

ば、原稿のタイプ、記憶容量、画像原稿の部分解析等のパラメータを用いることで、よりブロックセレクション処理の精度が向上するものである。

【0002】

【従来の技術】 米国特許出願第07/813,012号には、文字認識装置及びその方法、また、米国特許出願第08/171,720号には、読取原稿のテキスト/非テキスト選択装置及びその方法が記載されている。従来の米国特許出願第07/813,012号、08/171,720号に記載されているようなブロックセレクション処理は、原稿から得られる複数種類の画像データのタイプを区別するために、1ページ単位で自動的に画像原稿の解析を行うページ解析システムが用いられていた。

【0003】 上述のブロックセレクション処理の処理単位は、光学的文字認識(OCR)、データ圧縮、データ等で得られるイメージを、以下説明される処理によって複数のタイプに区別される単位が処理単位として用いられる。例えば、OCRによって認識されるテキストデータと、認識されない図形データとしてのイメージが、その処理単位である。その結果、ユーザによる操作を必要とすることなく異なるタイプの画像データが入力可能であり、また、自動的に処理が行われる。

【0004】 以下、ブロックセレクション処理の一手順を図1~3を参照して説明する。図1は1ページの原稿の代表的な例を示す図である。原稿101は、2つの段落から構成されており、タイトル102、水平線104、複数のテキストデータの行からなるテキスト領域105~107、非テキストであるグラフィックを含むハーフトーン図形領域108、テキストデータを含む表110、枠116、キャプション126を持つハーフトーン図形領域121、キャプション137を持つ図形領域132、135からなる。

【0005】 米国特許出願第07/873,012号、08/171,720号に記載されるブロックセレクション処理によれば、原稿101は構成される各領域は、画像データのタイプ毎に分類される。そして、ページ単位の原稿に関する、ブロックセレクション処理により、図2に示すような、階層ツリー構造が生成される。図2に示すように、階層ツリー構造200は、画像データを複数のブロックに分割される各ブロックが1つのノードである複数のノードから構成されている。各ノードは、ブロックセレクション処理によって得られる各ブロックの特徴を示す特徴データが付加されている。例えば、特徴データとしては、ブロック位置データ、サイズデータ、属性データ(例えば、テキスト、絵、表等の図形タイプ)、副属性データ、子ノードあるいは親ノードポイントがある。

【0006】 また、子ノードあるいは子孫ノードが、階層が上のブロックの下に存在する。子ノードは、階層ツ

7

リー構造の親ノードから分枝するノードとして示される。例えば、図2において、子ノードは、親ノードあるいはルートノード211から分枝するノード211と同レベルのノードが子ノードとなる。また、テキストブロックとなるノードには、その処理方向と処理順序（図に「RO＝」が示す番号）を示す特徴データが付加される。

【0007】例えば、階層ツリー構造200に示されるような階層ツリー構造が生成されたとなると、それはメモリに格納される。また、原稿より画像データを要求する処理において、メモリより階層ツリー構造が検索されブロック化された画像データに対応するノードの特徴データが生成される。また、原稿101に対する画像データを処理し、更に原稿101の簡単な構成を生成し、その結果を表示するために階層ツリー構造が用いられる。例えば、図3に示されるようなディスプレイ300に、簡単な構成であるブロックテンプレート301が表示される。また、ブロックテンプレート301は、図2の階層ツリー構造で示された特徴データに基づいて生成された表示される。

【0008】1ページの原稿に対応するブロックテンプレートは、階層ツリー構造ばかりでなく、階層ツリー構造の各ノードの特徴データに直接反映する。各ノードの特徴データは、ブロックの外枠、位置、サイズを生成することで利用される。また、各ノードの特徴データは、ブロックに含まれる画像データのタイプ、可能な割当、処理順、処理方向を認識するために用いられる。

【0009】例えば、図3に示されるブロックテンプレート301において、テキストブロック302～305、309は、それぞれ図2のノード202～205、209に対応している。また、それぞれのテキストブロックには、各テキストブロックの処理順、処理方向である特徴データを含んでいる。尚、図のT1～T14はそれぞれがテキスト1～テキスト14を示しており、また、P1CT、1～P1CT、4は図形1～図形4を示している。

【0010】

【発明が解決しようとする課題】しかしながら、上述のブロックセレクション処理は、以下に説明される不慮の処理に遭遇すると以下のような問題点が発生していた。例えば、精度の悪いOCRによって得られるテキストデータにおけるブロックセレクション処理は、エラーを含む処理結果が生じてしまっていた。また、線の結合（90度に交差する線、また枠でない線の結合）処理において、その線の抽出に際し、OCRの精度が悪いと正確な位置関係の線を抽出することができないために、エラーを含む処理結果が生じてしまうという問題点があった。更に、歪んだ原稿あるいは外部装置から入力される歪んで画像データのように、ブロックセレクション処理でエラーを生じる原稿あるいは画像データを事前に解析し、

8

そのエラーを抑止するようなアプリケーションは存在しなかった。

【0011】本発明は上記の問題点に鑑みてなされたものであり、ブロックセレクション処理の処理対象である原稿のデータ量と歪みを監視し、監視の結果に基づいて処理の実行の有無を指示することで、精度の良いブロックセレクション処理を実行できるページ解析システムを提供することを目的としている。

【0012】

【課題を解決するための手段】上記の目的を達成するための本発明による出力方法は以下の構成を備える。即ち、複数ページの画像データを解析するためのページ解析システムにおいて、該システムは、ブロックセレクション処理プログラムに従う処理を実行するために、複数ページの原稿の1ページ分の画像データを読み込み格納する第1記憶領域と、第1記憶領域に格納された画像データに対して実行されたブロックセレクション処理の処理結果を格納する第2記憶領域とを備え、該システムにおける記憶容量不足の指示の出力を行う出力方法であって、前記第1記憶領域に、1ページ分の入力画像データを格納する記憶領域と前記入力画像データのブロックセレクション処理を実行するための記憶領域を割り当てる第1割当工程と、前記第2記憶領域に、ブロックセレクション処理の複数の処理結果を格納するための記憶領域を割り当てる第2割当工程と、前記複数ページの原稿の解析されるべき1ページ分の画像データが、前記第1記憶領域に割り当てられた記憶領域の記憶容量を越える場合に第1エラーコードを出力し、前記第2記憶領域に格納されたブロックセレクション処理の処理結果を利用するための記憶領域を増やすことができない場合に第2エラーコードを出力する出力工程とを備える。

【0013】また、好ましくは、前記ブロックセレクション処理の処理結果が格納される前記第2記憶領域は、階層ツリー構造を含む。上記の目的を達成するための本発明による選択方法は以下の構成を備える。即ち、原稿1ページ分の画像データを解析するページ解析システムにおいて、前記原稿の少なくとも1つの領域を選択する選択方法であって、原稿1ページ分の画像データを入力する入力工程と、前記画像データを用いて前記原稿ページを表示する表示工程と、座標指定手段を用いて、解析すべき前記原稿ページの少なくとも1つの領域の座標を指定する指定工程と、前記指定工程で指定された前記原稿ページの少なくとも1つの領域内の画像データを抽出する抽出工程と、前記少なくとも1つの領域内の画像データを解析するために前記少なくとも1つの領域に対してブロックセレクション処理を実行する実行工程とを備える。

【0014】上記の目的を達成するための本発明による指定方法は以下の構成を備える。即ち、原稿1ページ分の画像データを解析するページ解析システムにおいて、

9

前記原稿を解析するために少なくとも1つの方向を指定する指定方法であって、原稿1ページ分の画像データを入力する入力工程と、前記画像データの方向情報を得る獲得工程と、前記方向情報が獲得されたか否かを判定する判定工程と、前記原稿を解析する解析工程を備え、前記方向情報が獲得された場合、前記解析工程は該方向情報によって指定される方向を用いて解析し、前記方向情報が獲得されない場合、前記解析工程は予め設定されている所定方向を用いて解析する。

【0015】また、好ましくは、前記所定方向は、水平方向と垂直方向が混合された方向である。上記の目的を達成するための本発明による獲得方法は以下の構成を備える。即ち、原稿1ページ分の画像データを解析し、ブロックセレクション処理によって前記された1つのルートノードと該ルートノードの子孫ノードを含む階層ツリー構造を生成するページ解析システムにおいて、前記原稿ページの傾き角度を獲得する獲得方法であって、前記原稿ページの画像データを入力する入力工程と、前記画像データに対し、前記画像データの傾き角度の算出を含むブロックセレクション処理を開始する開始工程と、前記算出された傾き角度と所定最大傾き角度を比較する比較工程と、前記算出された傾き角度が前記所定最大傾き角度よりも大きい場合、前記傾き角度が前記所定最大傾き角度を越えない場合、前記画像データのブロックセレクション処理を完了する完了工程と、前記傾き角度が前記所定最大傾き角度を越える場合、ブロックセレクション処理を中断する中断工程とを備える。

【0016】また、好ましくは、前記傾き角度は、前記階層ツリー構造のルートノードに格納される。上記の目的を達成するための本発明による出力装置は以下の構成を備える。即ち、複数のページの画像データを解析するためのページ解析システムにおいて、該システムは、ブロックセレクション処理プログラムに従う処理を実行するために、複数のページの原稿の1ページ分の画像データを読み込み格納する第1記憶領域と、第1記憶領域に格納された画像データに対して実行されたブロックセレクション処理の処理結果を格納する第2記憶領域を備え、該システムにおける記憶容量不足の指示の出力を行う出力装置であって、前記第1記憶領域に、1ページ分の入力画像データを格納する記憶領域と前記入力画像データのブロックセレクション処理を実行するための記憶領域を割り当て、前記第2記憶領域において、ブロックセレクション処理の複数の処理結果を格納するための記憶領域を割り当てると、前記複数のページの原稿の解析されるべき1ページ分の画像データが、前記第1割り当て領域に割り当てられた記憶領域の記憶容量を越える場合に第1エラーコードを出力し、前記第2記憶領域に格納されたブロックセレクション処理の処理結果を利用するための記憶領域を増やすことができない場合に第2エラー

10

コードを出力する出力手段とを備える。

【0017】また、好ましくは、前記ブロックセレクション処理の処理結果が格納される前記第2記憶領域は、階層ツリー構造を含む。上記の目的を達成するための本発明による選択装置は以下の構成を備える。即ち、原稿1ページ分の画像データを解析するページ解析システムにおいて、前記原稿ページの少なくとも1つの領域を選択する選択装置であって、原稿1ページ分の画像データを入力する入力手段と、前記画像データを用いて前記原稿ページを表示する表示手段と、解析すべき前記原稿ページの少なくとも1つの領域の座標を指定する座標指定手段と、前記座標指定手段で指定された前記原稿ページの少なくとも1つの領域内の画像データを抽出する画像データ抽出手段と、前記領域内の画像データを解析するために前記少なくとも1つの領域に対してブロックセレクション処理を実行する実行手段とを備える。

【0018】上記の目的を達成するための本発明による指定装置は以下の構成を備える。即ち、原稿1ページ分の画像データを解析するページ解析システムにおいて、前記原稿ページを解析するために少なくとも1つの方向を指定する指定装置であって、原稿1ページ分の画像データを入力する入力手段と、前記画像データの方向情報を得る獲得手段と、前記方向情報によって指定された方向を用いて解析し、前記方向情報が獲得されない場合、前記解析手段は該方向情報によって指定された方向を用いて解析し、前記方向情報が獲得されない場合、前記解析手段は予め設定されている所定方向を用いて解析する。

【0019】また、好ましくは、前記所定方向は、水平方向と垂直方向が混合された方向である。上記の目的を達成するための本発明による獲得装置は以下の構成を備える。即ち、原稿1ページ分の画像データを解析し、ブロックセレクション処理によって前記原稿ページの1つのルートノードと該ルートノードの少なくとも1つの子孫ノードを含む階層ツリー構造を生成するページ解析システムにおいて、前記原稿ページの傾き角度を獲得する獲得装置であって、前記原稿ページの画像データを入力する入力手段と、算出された傾き角度と所定最大傾き角度を比較する比較手段と、前記算出された傾き角度が前記所定最大傾き角度よりも大きい場合、前記傾き角度が前記所定最大傾き角度を越えない場合、前記画像データのブロックセレクション処理を実行するブロックセレクション処理手段とを備え、前記ブロックセレクション処理手段は、ブロックセレクション処理を開始し、算出された傾き角度と所定最大傾き角度を比較するように前記比較手段を用い、前記算出された傾き角度が前記所定最大傾き角度を越えない場合、ブロックセレクション処理を完了させる。

【0020】また、好ましくは、前記傾き角度は、前記階層ツリー構造のルートノードに格納される。上記の目



11

的を達成するための本発明による指示方法は以下の構成を備える。即ち、複数ページの原稿の画像データを解析するためのページ解析システムにおいて、ブロックセレクション処理プログラムによって解析されたページの階層ツリー構造が生成され、前記階層ツリー構造は1つのルートノードと該ルートノードの少なくとも1つの子孫ノードを含み、前記複数ページの原稿1ページの画像データの合成領域を指示する指示方法であって、前記複数ページの1ページ分の画像データを入力する入力工程と、前記画像データを識別し、同じタイプの画像データ

を有するブロックを認識へけるようにブロックセレクション処理を準備する準備工程と、同じタイプの画像データのブロックを他に持つ少なくとも1つのブロックを認識する認識工程と、前記同じタイプの画像データを持つ複数のブロックを1つの合成ブロックに結合する結合工程と、前記合成ブロックを、前記複数のブロックに対応する子孫ノードを持つルートノードとして前記階層ツリー構造に定時する定時工程とを備える。

【0021】上記の目的を達成するための本発明による指示装置は以下の構成を備える。即ち、複数ページの原稿の画像データを解析するためのページ解析システムにおいて、前記原稿ページを解析するための階層ツリー構造と、前記階層ツリー構造は1つのルートノードと該ルートノードの少なくとも1つの子孫ノードを含み、前記複数ページの原稿1ページの画像データの合成領域を指示する指示装置であって、前記複数ページの1ページ分の画像データを入力する入力手段と、前記画像データを識別し、同じタイプの画像データを有するブロックへけるブロックセレクション処理手段と、同じタイプの画像データのブロックを他に持つ少なくとも1つのブロックを認識する認識手段と、前記同じタイプの画像データを持つ複数のブロックを1つの合成ブロックに結合する結合手段と、前記合成ブロックを、前記複数のブロックに対応する子孫ノードを持つルートノードとして前記階層ツリー構造を変更する変更手段とを備える。

【0022】上記の目的を達成するための本発明による認識方法は以下の構成を備える。即ち、1ページ分の原稿の画像データのブロックテンプレートに従って階層ツリー構造を生成するページ解析システムにおいて、前記階層ツリー構造は複数のノードを持ち、各ノードは原稿画像のブロックテンプレートにおける表されるブロック毎の画像データに対応しており、また、各ノードには前記原稿画像のブロックテンプレートの特徴を定義する特徴データが含まれ、該システムにおいて図形タイプの指示及び認識方法であって、複数ページの原稿の1ページ分の画像データを入力する入力工程と、前記画像データを識別し同じ画像タイプのブロック毎に分割するブロックセレクション処理を実行する実行工程と、図形情報を含む画像データのブロックを認識する認識工程と、前記ブロックに含まれる図形情報のタイプを判定する判定工

12

程と、前記判定工程は、図形情報がハーフトーン、線画、折れ線、未知のいずれかを判定し、前記判定工程で判定される図形のタイプを、前記ブロックに対応する前記階層ツリー構造のノードに格納する格納工程とを備える。

【0023】上記の目的を達成するための本発明による認識装置は以下の構成を備える。即ち、1ページ分の原稿の画像データのブロックテンプレートに従って階層ツリー構造を生成するページ解析システムにおいて、前記階層ツリー構造は複数のノードを持ち、各ノードは原稿画像のブロックテンプレートにおける表されるブロック毎の画像データに対応しており、また、各ノードには前記原稿画像のブロックテンプレートの特徴を定義する特徴データが含まれ、該システムにおいて図形タイプの指示及び認識装置であって、複数ページの原稿の1ページ分の画像データを入力する入力手段と、前記画像データを識別し同じ画像タイプのブロック毎に分割し、分割されたブロックを認識するブロックセレクション処理と、図形情報を含む画像データのブロックを認識する認識手段と、前記ブロックに含まれる図形情報のタイプを判定する判定手段と、前記判定手段は、図形情報がハーフトーン、線画、折れ線、未知のいずれかを判定し、前記判定手段で判定される図形のタイプを、前記ブロックに対応する前記階層ツリー構造のノードに格納する格納手段とを備える。

【0024】

【発明の実施の形態】以下、実施の形態の概要を説明する。本実施の形態に従うページ解析システムは、複数ページの原稿の画像データを解析する。上述のシステムは、複数ページの原稿の1ページ分の画像データを第1の記憶領域に読み込み格納する。そして、第1の記憶領域に格納された原稿1ページ分の画像データに対してブロックセレクション処理が実行される。

【0025】また、第1の記憶領域に格納された1ページ分の画像データに対して実行されたブロックセレクション処理の処理結果が、第2の記憶領域に読み込まれ格納される。そして、第1の記憶領域（ブロックセレクション処理の処理対象となる1ページ分の画像データを読み込み格納する領域）と、第2の記憶領域（第1の記憶領域に記憶される画像データに対して実行されたブロックセレクション処理によって得られる複数の処理結果を読み込み格納する領域）に対し、以下に説明される場合に、各記憶領域の記憶容量不足を示すエラーコードを出力する。

【0026】第1の記憶領域で生じる記憶容量不足は、複数ページの原稿の1ページ分を超える画像データが読み込まれた場合に、第1の記憶領域の記憶容量不足を示す第1エラーコードが出力される。また、第2の記憶領域で生じる記憶容量不足は、第2の記憶領域に読み込み格納可能なブロックセレクション処理の有効な処理結果

13

を記憶する記憶領域が存在しない場合、第2の記憶領域の記憶容量不足を示す第2エラーコードが出力される。

【0027】また、他の実施の形態によれば、ページ解析システムは、原稿1ページ分の画像データを解析し、少なくとも、その原稿ページの一部分の領域が解析対象として選択される。原稿ページは、画像データとして処理され表示される。その際、座標指定装置によって、少なくとも解析対象である原稿の一部分の領域に座標が指定される。そして、選択された原稿ページの一部分の領域を含む原稿ページの画像データが抽出され、少なくともその抽出された領域の画像データを解析するために、ブロックセレクション処理が実行される。

【0028】また、他の実施の形態によれば、ページ解析システムは、原稿1ページ分の画像データに対し、少なくとも1つの方向を解析する。原稿1ページ分の画像データが入力されると、その画像データの方向情報が獲得される。そして、その方向情報が獲得されたか否かを判定する。判定の結果、方向情報が獲得された場合、その獲得された方向情報によって方向が選択され、その選択された方向に従って画像データが解析される。一方、方向情報が獲得されない場合、予め設定されている所定方向に従って画像データが解析される。

【0029】また、他の実施の形態によれば、ページ解析システムは解析対象として原稿1ページ分の画像データを入力し、その原稿ページの傾き角度の算出を含むブロックセレクション処理の実行を開始する。そして、算出された傾き角度と予め設定されている所定の傾き角度を比較し、算出された傾き角度が所定の傾き角度よりも大きいかなかを判定する。算出された傾き角度が所定の傾き角度よりも小さい場合、ブロックセレクション処理が完了する。一方、算出された傾き角度が所定の傾き角度よりも大きい場合、ブロックセレクション処理は中断される。

【0030】また、他の実施の形態によれば、ページ解析システムは、複数ページの原稿の1ページ分の画像データを入力し、その画像データを同じタイプの画像データのブロックに分割し、各ブロックの認識を行う。そして、同じタイプの画像データのブロックを持つ複数のブロックの少なくとも1つのブロックに対して、そのブロックの認識を行う。

【0031】尚、本実施の形態によれば、同じタイプの画像データのブロックは合成ブロックとして結合される。例えば、合成ブロックとしては、キャプションを含む図形等が合成ブロックとなる。また、合成ブロックは、階層ツリー構造において、同じタイプの画像データのブロックとして構成されるノードを下方に持つルートノードとして構成される。

【0032】また、他の実施の形態によれば、ページ解析システムは、複数ページの原稿の1ページ分の画像データを入力し、その画像データを同じタイプの画像データ

14

のブロックに分割し、各ブロックの認識を行う。そして、図形情報を含む画像データのブロックの認識を行う。尚、本発明によれば、図形情報を持つ画像データのブロックとしては、ハーフトーン図形、線画、折れ線、それ以外のタイプの図形がそのブロックの図形情報となる。そして、その図形と認識されたブロックは、そのブロックに対応する階層ツリー構造のノードに図形情報が付加された形で、そのノードが構成される。

【0033】尚、これらの本発明の特徴及び有利な点は、以下の添付の図面に従う好適な実施の形態の詳細な説明を参照することにより、更に容易に理解されるであろう。図4は本発明を適用する実施の形態の外観図である。以下に図4の構成を説明する。410はコンピュータ装置であり、Macintosh、IBM PCまたはMicrosoft社製Windowsのようなウィンドウズ環境を備えたPC互換コンピュータからなる。コンピュータ装置410には、カラーモニター等のディスプレイ412、ユーザコマンド等を入力するためのキーボード413、そして、ディスプレイ412上に表示される事象を指示し操作するためのマウス等のポインティングデバイス414が接続されている。

【0034】コンピュータ装置410は、大容量記憶装置として、イメージファイルを含むデータファイル記憶するためのディスク411を備えている。ディスク411は、圧縮、あるいは無圧縮のイメージファイルや、本実施の形態で説明されるブロックセレクション処理を実行するためのためのアプリケーションプログラムを含む情報処理プログラムを記憶している。また、ディスク411には、1ページ分の画像原稿のそれぞれがブロックセレクション処理によって表現される階層ツリー構造の各データも記憶している。

【0035】複数ページの画像原稿は、スキャナ416によってスキャンされる。スキャナ416によって、スキャンされた各ページを構成する文書あるいはイメージは、ビットマップ画像データとしてコンピュータ装置410へ展開される。また、画像データは、ネットワークインターフェース424またはファクシミリ/モデムインターフェース426を介して、他の外部装置からコンピュータ装置410へ入力してもよい。

【0036】プリンタ418は、コンピュータ装置410によって処理された画像原稿の出力のために備えられている。尚、図4では、プログラム可能な汎用コンピュータ装置としてその外観を示したが、専用、あるいは独立したコンピュータ装置、または他のタイプのデータ処理装置であっても本発明の実施の形態に適用できることは言うまでもない。

【0037】図5は実施の形態のコンピュータ装置410の詳細な内部構造を示すブロック図である。以下に図5の構成を説明する。図中、コンピュータ装置410は、コンピュータバス421によって、中央処理ユニッ

15

ト(CPU)520を含む各ユニットが相互に接続されている。コンピュータバス421に接続される各ユニットは、スキャニングファース522、プリンティングファース523、ネットワークインタフェース524、ファックス/モデムインタフェース526、ディスプレイインターフェース27、キーボードインタフェース528、マウスインタフェース529、メインメモリ530、そしてディスク411からなっている。

【0038】メインメモリ530は、コンピュータバス521に接続され、RAMを備えている。RAMは、CPU520で実行される各種処理(例えば、ブロックセ  
レクション処理)を実行するための作業領域である。また、メインメモリ530では、CPU520によってディスク411に記憶された各種処理を実行するためのプログラムが読み込まれ、読み込まれた各種プログラムによって各種処理が実行される。

【0039】ユーザの操作によって、画像データを処理するためのアプリケーションプログラムが読み込まれ、その処理が実行される。例えば、ウィンドウズ版WordPerfectのようなデスクトップ処理プログラムが操作する  
ことで画像データが処理され、また、ブロックセレクション処理前後の画像原稿の処理状況を表示する。同様に、ページ解析プログラムが、ブロックセレクション  
処理を行うことで処理対象の1ページ分の画像原稿に複数種類の画像データの属性を入力する。そして、その1ページ分の画像原稿の処理結果がウィンドウ環境を介してディスプレイに表示される。

【0040】図6A~6Dは本実施の形態のブロックセレクション処理によるページ解析の方法を示すフローチャートである。尚、図6A~6Dで説明される各処理のステップは、ディスク411によって読み込まれるコンピュータプログラムに従って、CPU520によって実行される。

【0041】ステップS601では、画素画像データがシステムへ入力されディスク411へ格納される。画像データは画像を画素単位で表現している。好ましくは、画素データは2値画素データ、つまり白黒画像データである。しかし、画像データは、各画素が複数のグレースケールレベルの1つで表現される中間調画像データであっても良いし、各画素が、その色を符号化するマルチビットワードで表現されるカラー画像データであっても良い。またの場合、または画素データが2値画素データでない他のいずれの場合でもディスク411へ格納する前に、非2値画素データを2値画素データへ変換するための閾値処理が行われる。

【0042】ステップS602で、ブロックセレクション処理プログラムが、そのプログラムと変数によって生成される複数の内部変数がセットされることによって初期化される。ステップS603で、ユーザによって、ブロックセレクション処理によって解析される1ページ分

16

の原稿が入力されたか否かが確認される。ブロックセレクション処理による解析が原稿ページの一部の領域だけしかない場合、ステップS604で、その解析する領域の調整を行う。

【0043】ここで、ステップS604において、解析される原稿ページの少なくとも一部分の領域の選択方法の概略について、図7A、図7Bを用いて説明する。図7Aは本実施の形態の原稿1ページの一部分の領域の解析処理の詳細な処理フローを示すフローチャートである。ステップS701で、原稿1ページ分の画像データを入力する。例えば、図7Bに示されるような原稿ページ750を入力すると、入力される原稿ページ750の画像データが読み取られディスク411に格納される。ステップS702で、入力された原稿ページの画像データがディスク411から検索され、その検索結果がユーザへと表示される。

【0044】ステップS703で、入力された画像データの一部分の領域を解析するのx y座標の入力がユーザによって要求される。例えば、原稿ページ750のテキスト領域751を解析するために、ユーザは座標(x1, y1)、(x2, y2)、(x3, y3)、(x4, y4)を入力する。ステップS705で、ユーザによるx y座標の入力後、ページ解析プログラムは、ステップS703で入力した座標が原稿ページの境界線内に含まれているか否かを判定する。例えば、図7Bに示されるような原稿ページ750に対し、その原稿ページ750の境界線内に入力された座標が含まれ  
るか否かがステップS705において判定される。原稿ページの境界線内に入力された座標が含まない場合は、ステップS703に戻る。そして、原稿ページの境界線内に含む座標が得られるまでユーザは座標の入力を繰り返す。あるいはユーザが解析対象にすべての原稿ページが入力されたことを指示するまで座標の入力を繰り返す。一方、原稿ページの境界線内に入力した座標が含ま場合は、ステップS706に進む。

【0045】ステップS706で、入力した座標によって定義される原稿ページの一部の領域の画像データの抽出を行う。抽出された画像データは、ステップS710のブロックセレクション処理における処理対象となる。ここで、再びステップS605に戻る。1ページ分の原稿の入力処理が終了するか、あるいはステップS604で原稿ページの一部の領域の選択が示され、ステップS605に進む。

【0046】ステップS605で、1ページ分の原稿あるいはステップS604選択された原稿ページの一部の領域に対し、ブロックセレクション処理を実行するための記憶領域を確保する。また、原稿ページの解析によって得られる階層ツリー構造の結果を格納するための記憶領域を確保する。ここで、ステップS605における記憶領域の確保する方法の概略について、図8を用いて

17

説明する。

【0047】図8は本実施の形態のブロックセレクション処理のための記憶領域の確保の方法と、階層ツリー構造を算出する処理のための記憶領域の確保の方法を説明するための図である。例えば、原稿ページ800は共通原稿ページ記憶領域が入力されている原稿ページである。共通原稿ページ記憶領域は、CPU520の制御のための作業領域であり、その制御とは、原稿ページをスキャンして得る画像データを記憶する記憶領域（図の記憶領域801a）を読み込む際の制御と、ブロックセ

レクション処理を実行するための作業領域（図の記憶領域801b）を確保するための制御である。

【0048】尚、CPU520が生成するこの記憶領域は、原稿ページの画像データを読み込むための記憶領域ばかりでなく、ブロックセレクション処理の中間処理結果を読み込むための記憶領域になっている。記憶領域803は、複数ページからなる記憶領域で構成されている。図に示されるように、804、805、806の各ページには、原稿ページ800のブロックセレクション処理によって得られる複数の階層ツリー構造が格納される。

【0049】尚、各ページの階層ツリー構造は異なるタイプの構造（あるページは単純な構造、また、あるページは複雑な構造）から構成されているため、各ツリー構造を記憶するための記憶容量は異なる。尚、初期化の際、共通原稿ページ記憶領域801と階層ツリー構造記憶領域803は、ブロックセレクション処理プログラムに設定されている変数に基づいて、充分な所定記憶容量の記憶領域が確保される。但し、ブロックセレクション処理プログラムによって確保される共通原稿ページ記憶領域801と階層ツリー構造記憶領域803の記憶領域は、その記憶容量はユーザによって変更しても良い。また、10ページだけといった所定数のページ数に基づいて記憶領域を確保したり、所定の変数に基づいて記憶領域を確保しても良い。

【0050】ステップS605における記憶領域が確保されると、ステップS607で、入力された原稿ページのタイプがユーザによって判定される。例えば、ステップS601でスキャンされた原稿ページの画像データが、英語原稿ページのような横書き原稿ページ、即ち水平であるか、あるいは日本語原稿ページのような縦書き原稿ページ、即ち垂直であるかが判定される。また、原稿ページが、縦書き横書きの両方から構成される場合は、その際は複数の方向を含む混合タイプとしての処理が必要となる。いずれの場合にしても、ユーザは、所定方向、垂直方向、水平方向、あるいは複数の方向を含む混合方向の4つのいずれかの方向を判定する。

【0051】尚、「混合」タイプの原稿ページは予め設定されたタイプであり、混合タイプの原稿ページが選択されても、ユーザはそれに伴う特別な処理は必要とされ

18

ない。ステップS608で、ユーザは入力した原稿ページのタイプが既に設定されているタイプであるかを判定する。入力した原稿ページのタイプが既に入力された原稿ページのタイプである場合、ステップS609で、そのタイプの情報が原稿ページのルートノードに格納される。また、原稿ページのタイプは、原稿ページに対して実行される処理に渡って、変数として用いられ、また、その変数は正確な処理結果を出力するための要素となっている。例えば、スキャンした原稿ページにおいて、テキストデータがOCRプログラムによって光学的に認識された場合、そのテキストデータの方向として水平方向がセットされる。

【0052】一方、ステップS608で、入力された原稿ページのタイプが既に設定されているタイプでない場合は、ステップS611で、その原稿ページのタイプを設定する。尚、本発明では、上述のステップS608で設定される原稿ページのタイプを「混合」タイプの原稿ページとして扱う。

【0053】そして、設定された原稿ページのタイプの情報が原稿ページのルートノードに格納される。また、原稿ページのタイプは、原稿ページに対して実行される処理に渡って、変数として用いられる。そして、原稿ページのタイプがセットされると、ステップS615からステップS635で説明されるブロックセレクション処理が実行される。

【0054】ブロックセレクション処理について簡単に説明する。ブロックセレクション処理は、2値画像の原稿ページにおいて、以下の異なるタイプのブロックに分割する。

(1) テキストデータとして、テキストデータを含む段落、タイトルテキストデータ、表の中のテキストデータ

(2) 線として、水平線、垂直線、ドットからなる線、斜線

(3) 互いに直交する線分

(4) ハーフトーン画像、連続調、グレースケール等

(5) 罫面

(6) 例えば、箱型の枠、枠である領域として認識されるもの（例えば、図1の枠116）

(7) 表（例えば、図1の表110）

(8) 枠や線画の中の空白領域から独立して存在する空白領域

更に、従来のブロックセレクション処理とページ分割は、横きのない原稿ページであることを前提に処理を行うブロックセレクション処理とページ分割であった。しかし、本発明によれば、後述するステップにおいて、原稿ページの横きを検知し、横きのある原稿ページ、横きのない原稿ページのいずれにおいてもそれぞれに対応する処理を実行すると、横きのある原稿ページに対する処理が保証されている。この保証の方法としては、原

19

稿ページに傾きがある場合、ステップS615からステップS635で説明される処理によって分割される各ブロックが長方形であるブロックは重ならない表示されてしまう。また、原稿ページに傾きがない場合は、それらの長方形であるブロックは重ならない。この事実をふまえて、原稿ページを処理する際に生じる長方形であるブロックの重なりの有無によって、原稿ページが傾いているか否かを知る。そして、原稿ページに傾きがある場合は、その原稿ページに対するブロックの抽出する際にその旨を示す情報を与えることで、傾きのある原稿ページに対する処理が保証される。

【0055】ステップS615からステップS635の処理の概要を説明する。ステップS615からステップS635で説明される処理によって、階層ツリー構造は原稿ページの画像データを表現するために用いられる。また、階層ツリー構造のルートノードは原稿ページの含むすべてのテキストブロック、非テキストブロックを表現し、それらのノードはルートノードの直接の子孫あるいはその子孫の子孫としてつながったノードの状態で表現される。

【0056】一方、原稿ページをブロックに分割するために、原稿ページの画像データの結合成分が検索され、その結合成分によってテキスト部と非テキスト部の結合成分に区別される。そして、テキスト部と非テキスト部の結合成分はおおまかにかつ簡単にグループ化される。例えば、いくつかの段落によって構成されるテキスト部の結合成分が集められてテキストブロックが生成される。また、例えば、一つの絵として存在する結合成分が集められることによって絵ブロックが生成される。

【0057】このより、ステップS615からステップS635の詳細な処理について説明する。ステップS615で、画像データが縮小された処理が実行される。ブロックセレクション処理の処理速度を上げるために画像データの縮小は必要であるが、光学的文字認識(OCR)に影響しないような画像データの縮小を行う。画像データの縮小方法としては、例えば、解像度が400dpiの原稿ページを4画面毎に1画面間引いて200dpiの解像度にする程度の間引きは、画像の結合性が保証されることが経験的に証明されている。また、他の間引き方法としては、異なる解像度を持つ画像を用いる方法である。例えば、解像度500dpiの画像と、解像度750dpiの画像を生成し、その2つの結果を用いることで、正確でかつ高速に画像の間引きが実現されることが経験的に知られている。

【0058】ステップS616で、画像画像データが解析された連結成分が検出され、その大きさ及び他の連結成分との相対的な位置に基づいて、検出された連結成分がテキストであるか非テキストであるかが区別される。連結成分は、米国特許出願第07/813,012号、08/171,720号に記載されているように、白画

20

素によって完全に囲まれた黒画素のグループである。連結成分の検出は、例えば、図9Aに示すように連結成分902の輪郭901を検索することによって、連結成分の検出が開始される。この検出は、図9Aに示されるような所定8方向の各方向に存在する連結成分の輪郭の画素成分を検出することで行う。輪郭901に囲まれる連結成分901を構成する内側の各画素は、後述するステップS618で区別されるテキストあるいは非テキストに基づいて、その画素が飛ばされて解析される。そして、各連結成分は、各連結成分の周りに構成することができる最小の長方形を使うことによって、各連結成分を長方形化する。

【0059】ステップS617で、連結成分のサイズに基づいて、各連結成分はおおまかにテキスト部と非テキスト部に区別される。尚、一般的には、非テキスト部の方がテキスト部よりも大きくなる。ステップS617の区別で誤った区別がされてしまった場合、以下に説明されるステップで補正される。その補正方法としては、テキストと非テキストを区別するために、各連結成分のサイズと所定のサイズの閾値を比較することで、その区別を行うための判定が行われる。また、その所定のサイズの閾値は、各連結成分に基づいて得られる適応的に決定された閾値である。

【0060】詳細に説明すると、ステップS617のテキスト部と非テキスト部の区別は2つのステップで行われる。最初のステップで、極端に大きい連結成分は非テキスト成分として区別される。第2のステップで、区別されていない各連結成分を囲む長方形の高さと幅の平均値を算出する。算出した平均値に基づいて、適応的なサイズ(即ち、テキストを囲む長方形のサイズである垂直方向のテキストのサイズあるいは水平方向のテキストのサイズ)が算出され、その算出されたサイズが上述のテキスト/非テキストの区別のための閾値として用いられる。

【0061】ステップS617の処理によって得られたテキストと非テキストの連結成分それぞれは、階層ツリー構造のルートノードの直接の子孫のノードとして割り当てられる。テキスト連結成分と非テキスト連結成分の区別は、それぞれ同レベルのツリー構造のノードに対して行われる。また、その際、子孫レベルの連結成分の処理の過程で得られる統計的な値に従って、その区別は行われる。この区別は、より多くの階層ツリー構造がある場合に、より重要な処理となる。例えば、枠の外側と内側にサイズの一致しないテキストがある場合、枠の内側が区別が枠の外側の区別から先に他の処理は、異なる子孫レベルに移る処理となるため、同じ枠に対してその枠の内側の区別とその枠の外側の区別を別々に行うことになるからである。

【0062】ステップS618で、非テキスト連結成分は、それらが線(水平線、垂直線、ドット線あるいは斜

21

線)、折れ線、絵、線画、枠、表あるいは未知(即ち、どれにも属さない)のいずれかに判定される。また、階層ツリー構造のルートノードの子孫のその下の子孫である表、枠、線画、絵に含まれる非テキスト連結成分は、上述のステップS 617、ステップS 618を繰り返すことによって区別され判定される。そのため、各連結成分とその連結成分の種類に対して階層ツリー構造が展開される。しかしながら、枠や絵に含まれる連結成分の区別は、8方向の検索(図9Aを参照)における黒画素の検索は実行されない。そのかわりに、4方向の検索による白画素の検索が実行される。その検索方向は図9Bに示されるような方向であり、例えば、画素枠904で囲まれる白輪郭903に関して、4方向の検索を行う。この白輪郭の4方向の検索を行うのは、囲まれ輪郭の内側にある連結成分の検索と区別を行うのに役立つからである。

【0063】ステップS 618における非テキスト連結成分の区別、つまり、線、折れ線、枠、表、絵、線画、未知のいずれかのタイプとして区別される手順の好適な処理フローを図10に示す。図10に示されるような処理が、各非テキスト連結成分に対して実行される。また、非テキスト連結成分の各タイプの区別を実行する処理が、図10に示されるようなシーケンシャルな処理で実行される。

【0064】例えば、図10において、早い処理の段階で処理が実行される非テキスト連結成分は、ハーフトン画像と認識される。このように、非テキスト連結成分毎に処理が進む。また、図10で実行される処理の処理対象である非テキスト連結成分は、図10の各ステップの処理の実行の有無を基準となる閾値を比較することで決定する。その基準となる閾値は、数学的に公式化された計算によって、直接各連結成分を囲む長方形のサイズと幅に基づいて算出される。

【0065】図10の処理フローについて説明する。ステップS 1001で、非テキスト連結成分は、ハーフトン画像(あるいは連続画像)であるかが判定される。ハーフトン画像であるかどうかの判定は、その非テキスト連結成分の画像エリアに関し、白と黒画素のそれぞれの連結成分の比較と黒画素密度に基づいて行われる。その画像エリアに関し、黒画素密度が白画素成分よりも長い場合、また、黒画素密度が高い、あるいはその画像エリアに多くのノイズのような連結成分がある場合は、その非テキスト連結成分はハーフトン画像連結成分と判定される。そして、ステップS 1002に進み、その非テキスト連結成分に対応するノードに補助属性として「ハーフトン」が格納される。

【0066】一方、非テキスト連結成分がハーフトン画像連結成分と判定されない場合、ステップS 1003に進み、その非テキスト連結成分が枠であるかが判定される。枠であるかどうかの判定は、非テキスト連結

22

成分の高さがテキストサイズよりも長く、その非テキスト連結成分が、平行である水平な2本の境界線と平行である垂直な2本の境界線が、長方形を構成する境界線と一致する場合、その非テキスト連結成分は枠と判定される。また、非テキスト連結成分の黒画素密度がテキストデータ構成される段落の密度と近似あるいはそれよりも少ない場合(ステップS 10001におきて、非常に高い黒画素密度を持つ非テキスト連結成分と判定される非テキスト連結成分)、枠と判定される。非テキスト連結成分が枠と判定された場合、ステップS 1004に進み、その非テキスト連結成分に対応する階層ツリー構造のノードに補助属性として「枠」を格納する。

【0067】ステップS 1003で、非テキスト連結成分が枠と判定されなかった場合、ステップS 1005で非テキスト連結成分が水平線であるいは垂直線であるかが判定される。水平線あるいは垂直線であるかどうかの判定は、その非テキスト連結成分の幅と高さの割合が大きい場合や、また、階層ツリー構造の同じレベルのノードのテキスト連結成分の細さよりも非テキスト連結成分の細さの方が細い場合、その非テキスト連結成分は水平線と垂直線と判定される。更に、非テキスト連結成分に占める黒画素成分の密度がとても高い場合に、その非テキスト連結成分は水平線あるいは垂直線と判定される。水平線あるいは垂直線と判定される非テキスト連結成分の両側は真っ直ぐになっているべきであるが、その非テキスト連結成分の片側だけが真っ直ぐで、また非テキスト連結成分の細さがテキスト連結成分の細さと近似している場合は、下線である非テキスト連結成分と判定される。この場合、下線が引かれているテキスト連結成分と下線である非テキスト連結成分は互いに分割され、それぞれステップS 617とステップS 618による処理を改めて実行する。このようにして、非テキスト連結成分が水平線あるいは垂直線と判定された場合は、ステップS 1006に進み、その非テキスト連結成分に対応する階層ツリー構造のノードに補助属性として「水平線」あるいは「垂直線」を格納する。

【0068】ステップS 1005で、非テキスト連結成分が水平線あるいは垂直線と判定されない場合は、ステップS 1007で非テキスト連結成分が表であるかが判定される。表であるかどうかの判定は、非テキスト連結成分内の白輪郭が4方向による検索によって判定される。4方向の白輪郭が存在する場合、または、水平線と垂直線によって囲まれる白輪郭があるような表を構成する格子の中の白輪郭の配列である場合は、その非テキスト連結成分は表と判定される。このようにして、非テキスト連結成分が表と判定された場合は、ステップS 1008に進み、その非テキスト連結成分に対応する階層ツリー構造のノードに補助属性として「表」を格納する。また、表の内部にテキスト連結成分と非テキスト連結成分が認識された場合、その区別をするためにステップS

23

617とステップS618による処理を再び実行し、区別されたテキスト連結成分あるいは非的連結成分に対応する階層ツリー構造のノードを付加する。

【0069】ステップS1007で、非テキスト連結成分が表と判定された場合、ステップS1009で非テキスト連結成分が斜線であるか否かが判定される。斜線であるか否かの判定は、非テキスト連結成分の長さとし、その割合が大きい場合に、斜線と判定される。但し、その非テキスト連結成分の細さは、テキストサイズの細さよりも決して細くならない。また、非テキスト連結成分に占める黒画素密度が高密度で、そのエッジが斜めに並んでいる場合に斜線と判定される。このようにして、非テキスト連結成分が水平線あるいは垂直線と判定された場合は、ステップS1010に進み、その非テキスト連結成分に対応する階層ツリー構造のノードに補助属性として「斜線」を格納する。

【0070】尚、斜線の傾き角度は、既知の角度検出処理によって判定される。また、斜線の傾き角度は、後述する処理に用いるための属性情報として、階層ツリー構造の対応するノードに格納される。ステップS1009で、非テキスト連結成分が斜線と判定されない場合は、ステップS1011で非テキスト連結成分が線画あるいは折れ線であるか否かが判定される。線画あるいは折れ線であるか否かの判定は、非テキスト連結成分のサイズが前の処理で設定されているサイズよりも長い場合、例えば、ステップS1001で得られるハーフトーン画像のサイズよりも長い場合に、線画あるいは折れ線と判定される。また、線画はハーフトーン画像と比べて、白輪郭の連結成分が輪郭の内側における黒画素連結成分よりもかなり長い。加えて、黒画素密度が低い。このような条件を満たす場合、ステップS1012において、線画と折れ線を区別する処理が実行される。

【0071】図12Aは、折れ線と線画の違いを説明するための図である。図12Aにおいて、原稿ページ1050は、タイトル1051、テキスト領域1052、1053、絵1055からなる。テキスト領域1052、1053は、折れ線1054によって、絵1055から互いの分割されている。尚、一般的に折れ線は、図12C、図12Dに示す2つのグループからなる形状をもつものを折れ線と区別される。図12Cに示されるような折れ線は、改めて折れ線として選択され、図12Dに示されるような折れ線は、改めて枠として区別される（即ち、ステップS1003では枠として完全には選択されていない）。そして、それぞれの区別は、その構成要素の長さとし、互いに交差する角度（水平成分の鏡像、垂直成分の鏡像、対角線の鏡像でも許される）に基づいて行われる。

【0072】図12Aにおいて、折れ線1054は図12Cに示される折れ線に相当するので、この場合、ステップS1012からステップS1014に進む。このよ

24

うにして、非テキスト連結成分が折れ線と判定された場合は、ステップS1014に進み、その非テキスト連結成分に対応する階層ツリー構造のノードに補助属性として「折れ線」を格納する。

【0073】一方、非テキスト連結成分が図12Aに示される折れ線（「折れ線」の属性を付加するグループ）、あるいは図12Bに示される折れ線（「枠」の属性を付加するグループ）のどちらにも相当しない場合、ステップS1012からステップS1013に進み、非テキスト連結成分は線画と判定される。このようにして、非テキスト連結成分が線画と判定された場合は、ステップS1013に進み、その非テキスト連結成分に対応する階層ツリー構造のノードに補助属性として「線画」を格納する。

【0074】図12Bは、図12Aに示される様々な画像データが本発明のブロックセレクション処理による処理結果を示す図である。図12Aに示されるタイトル1051、テキスト領域1052、1053、絵1055は、図12Bにおいて、それぞれブロック1060、1061、1062、1063に対応する。また、図12Aに示される折れ線1054は、ブロック1063にブロック1061、1062が重なることで示されている。このように、それぞれのブロックの属性と調査結果は、階層ツリー構造に対応するノードに格納され、それに基づいて後述する画像データの抽出処理が実行される。

【0075】再び、図10の説明に戻る。ステップS1011で、非テキスト連結成分が線画あるいは折れ線と判定されない場合は、ステップS1015で非テキスト連結成分は未知のタイプとして判定される。このようにして、非テキスト連結成分が未知のタイプと判定された場合は、ステップS1015に進み、その非テキスト連結成分に対応する階層ツリー構造のノードに補助属性として「未知」を格納する。

【0076】更に、図6の説明に戻る。ステップS618による非テキスト連結成分の区別処理後、ステップS619に進み、ドットラインの検出が実行される。尚、一般的に連結成分が形成するドットラインの各ドットの検出は困難とされており、これらの連結成分は各ドットのサイズが同じであるテキスト連結成分とみなされる。

【0077】このようにドットラインの検出においては、水平成分あるいは垂直成分により近いサイズのドットサイズを持つテキスト連結成分が選択される。選択されたテキスト連結成分は各々の相対的な位置関係に基づいて異なるグループに区別される。区別されるそれぞれのグループは、サイズと付近のラインの囲みに基づくドットのラインが、垂直、水平、斜めの成分が正確に現れているか否かが確認される。それぞれ検出されるドットラインは、すべてのドットラインの要素に囲まれる輪郭を持つ連結成分で生成される。また、それぞれ検出され

25

るドットラインは、階層ツリー構造の対応するノードに、水平、垂直、斜めのいずれかのドットラインとしての属性が新しいユニットとして付加されて用いられる。

【0078】また、ステップS619では、折れ線であるドットラインの検出を行う。特に、グループ化された連結成分が水平なドットラインと垂直なドットラインが直角あるいは直角に近いかたちで交差している場合、そのグループ化された連結成分は、折れ線であるドットラインとして区別される。そして、検出される折れ線であるドットラインは、階層ツリー構造の対応するノードに、折れ線であるドットラインとしての属性が付加される。

【0079】ステップS620で、非テキスト連結成分のエッジに沿って不可視線が検索される。このような不可視線は横の間にある白線であり、この不可視線を検出することで、テキスト連結成分がグループ化されるべきかあるいは横から区別されるブロックにすべきかを判定するの役立つ。ステップS621で、ステップS607からステップS611の処理によって格納された原稿ページが「単一方向」である原稿ページのタイプに対し、その原稿ページの方角を検出する(図のステップS621a)。原稿ページの方角の検出は、原稿ページに存在する複数の連結成分において、その各連結成分の中でサイズが大きい連結成分のサイズに基づいて行う。例えば、水平方向である原稿ページの連結成分は、高さよりより同一の幅を持つことが多い。加えて、水平線と垂直線の数の差とサイズは、原稿ページの方角を判定するための付加的な要因になっている。そして、ステップS621aで検出される方向は階層ツリー構造のルートノードに格納される。なぜなら、単一方向である原稿ページは、グローバル変数となるからである。

【0080】ステップS622で、ステップS618で区別できなかった非テキスト連結成分(即ち、「未知」のタイプを持つ非テキスト連結成分)が解析される。そして、大きいフォントサイズのタイトルラインであるかどうか判定される。テキストあるいは未知のユニットを持つグループ化された「未知」のタイプの非テキスト連結成分によって、水平または垂直方向にタイトルラインが形成される。グループ化されたタイトルの方向は、グループ化されると想定される連結成分の間の距離と可能なタイトルの長さに基づいて判定される。

【0081】ステップS623で、テキストブロックがテキスト連結成分から形成される。ここでのステップでは、テキストあるいは非テキストブロックが形成される前には原稿ページの傾き検出も画像回転も行われることはない。この結果、画像回転に費やされる時間が節約され、原稿ページの傾き検出によって生じる不確かさ(文字認識処理(OCR)によって生じる不良)を避けることができる。更に、ある特殊な視覚効果を達成するために、同じ原稿ページのテキスト領域のいくつかを、

26

特別に斜めに編集することも可能である。このため、原稿ページに傾き角度がある場合でも傾き角度のない原稿ページと同様な効果得られるので、原稿ページを直立に補正するための原稿ページの傾きの検出は役に立たない。それゆえ、ステップS621では、原稿ページの傾き角度に関わりなく直接、テキスト及び非テキストブロックが形成される。

【0082】特に、それぞれのテキスト連結成分に対し、そのテキスト連結成分の近傍に存在する水平あるいは垂直である別のテキスト連結成分が検索され、その検索されるテキスト連結成分がテキストブロックとして集められる。この検索によってテキストブロックを集めるか否かの決定方法は、まず、このテキスト連結成分に対応する階層ツリー構造のノードと同じレベルにおいて、近傍にあるテキスト連結成分の水平あるいは垂直方向の隙間を統計的解析によって算出する。そして、算出された値と、所定の基準値である閾値とを比較することによって集めるか否かを決定する。それゆえ、各テキスト連結成分の接近の度合いが、原稿ページとその原稿ページ内で分割される領域に基づいて直接判定される。

【0083】ステップS624で、ステップS623の処理の際に集められてしまったテキストブロックに対し、そのテキストブロックに不可視線または不可視線が横切っている場合、そのテキストブロックは分割される。特に、ステップS620で検出される不可視線は、原稿ページに存在する極端に狭い横間の隙間によるものである。例えば、その隙間は図形領域の近くにテキスト横が横く場合が生じる。

【0084】また、テキストブロックと非テキストブロックの間にある不可視線の分割は実行されない。しかし、そのような不可視線に対しては、ステップS620において適切な分割が実行される。もちろん、物理的な分割も存在するが、その場合、水平成分、垂直成分、斜めの成分のいずれかを持つテキストブロックにおいては、ステップS624で集められたテキストブロックは分割される。

【0085】ステップS625で、ステップS607からステップS611の処理によって格納された各テキストブロックの方向が混合方向を持つ原稿ページであるか否かを判定する。特に、混合方向を持つ原稿ページタイプである場合、ステップS625aで、その原稿ページの各テキストブロックの方向が水平、垂直あるいはそれ以外の方向であるか否かが判定される。この方向(ステップS621aで判定された方向あるいは、ステップS609でユーザによって入力された方向)は、後述するステップにおいて、各テキストブロックのテキストラインを形成するために用いられる。更に、テキストブロックの方向は、近傍にある2つのテキストブロックが結合可能であるか否かを判定するために役立つ。例えば、その2つのテキストブロックの方向が一致しない場合、



27

通常はその2つのテキストブロックは結合されてしまうからである。

【0086】テキストブロックの方向の判定は、そのテキストブロックのサイズと、そのテキストブロックの中の連結成分のサイズとそのサイズに算出して得られるテキストブロックのサイズの算出量に基づいて行う。例えば、水平方向の連結成分を持つテキストブロックは一定の高さより幅を持っている。更に、水平方向と垂直方向の隙間の数とそのサイズは、テキストブロックの方向を判定するための付加要素として考慮される。

【0087】ステップS626で、各テキストブロックのテキストラインを形成する。例えば、テキストブロックの方向が水平である場合、各連結成分の水平方向の距離が接近して、また、垂直方向に位置する近傍の連結成分と所定角度で重なっている連結成分は水平方向のテキストラインとして集められる。同様に、テキストブロックの方向が垂直である場合、各連結成分の垂直方向の距離が接近して、また、水平方向に位置する近傍の連結成分と所定角度で重なっている連結成分は垂直方向のテキストラインとして集められる。テキストラインの構成は、真真正正な原稿ページばかりでなく、ある傾き角度を持つ原稿ページに対しても考慮する。

【0088】ステップS627で、原稿ページの傾き角度を検出する。原稿ページの傾きは、ステップS626で形成された各テキストラインの傾きに対し、最小2乗法を用いて算出する。原稿ページの傾きは、その原稿ページの各テキストラインの傾き角度に対し最小2乗法を用いて算出される平均の傾き角度である。ステップS627で算出された傾き角度は、その原稿ページの階層ツリー構造のルートノードにグローバル変数として格納される。ステップS607からステップS611の処理を通して格納される原稿ページのタイプと同様に、その原稿ページの傾き角度は、例えば、画像データの処理過程に必要な情報として光学式認識(OCR)処理あるいは他のタイプの処理が実行される際に、グローバル変数として用いられる。

【0089】ステップS628で、ステップS627で算出された傾き角度が所定最大傾き角度を超えるか否かを判定する。特に、ステップS628では、傾き角度はブロックセクション処理で設定される。また、傾き角度が所定最大傾き角度を超える場合、ステップS628で、原稿ページの傾きが大きすぎることをユーザに報知する信号としてのエラーコードが、ブロックセクション処理によって出力される。

【0090】ステップS629で、後処理が実行される。後処理は、簡潔で明瞭にブロックを表現するために実行される。また、後処理は、キャラクタ認識やデータ圧縮等の追加処理に適用させることもできる。更に、後処理において、テキストブロックは、他のテキストブロックあるいは非テキストブロックを結合し、また、他の

28

非テキストブロックあるいはテキストブロックも結合する。

【0091】後処理は、ステップS627で獲得される原稿ページの傾き角度に基づいて適用される。例えば、原稿ページの欄の隙間がほんのわずかに傾いている原稿ページは更に明瞭にされる。つまり、それらのブロック(欄)は積極的に結合される。一方、例えば、原稿ページの傾き角度が大きく、複数のブロックの長方形領域が重なっている場合、それらのブロックはより控えめに結合される。特に、原稿ページの傾き角度が小さい程、テキストブロックはそのテキストブロックの一部の領域に基づいて解析された欄情報に従って結合される。テキストブロックの結合は、基本的には同じ欄に含むテキストブロックであるかどうかに関係なく、つまり、例えば、それらのテキストブロックが互いに接近しているか、それらのテキストブロックの方向が一定であるか、それらのテキストブロックの結合が他のテキストブロックとの重なりを生じるか等によって結合の実行の有無を決定する。

【0092】一方、非テキストブロックの結合は非テキスト成分を含む主要な複数の非テキストブロックの部分に依存する。例えば、いくつかの図形は、いくつかの非連結成分によって形成される。これらのいくつかの非連結成分が同一の図形と判定されると、階層ツリー構造のダミーノードとして新たな合成領域が描画される。つまり、この合成領域はすべての非連結成分を包含するために生成される。更に、いくつかのテキストブロックが、同一図形として判定される場合、例えば、キャプションのようなテキストブロックの場合、それらのキャプションは合成領域として生成される。また、その合成領域は図形とテキスト領域を包含するダミーノードとして階層ツリー構造で用いられる。合成領域における処理の詳細は、後述のステップS634で説明する。

【0093】後処理が実行されると、テキストブロック、線ブロック、図形ブロック、枠ブロック、非テキストブロック、合成領域ブロック、表ブロックのいずれかのブロックに従う特定処理のための第1ブロック情報が選択される。本実施の形態によれば、特定処理は図形あるいは合成領域のブロックに対して必要とし、それ以外の他のブロックでは特定処理は行わない。何故なら、図形あるいは合成領域以外の他のブロックは、ブロックセクション処理によって必要とされる処理が完了するからである。図形と合成領域における特定処理について以下に説明する。図11は、処理対象のブロックが図形情報を持つ画像データとして認識された場合(ステップS632)に、実行される特定処理の処理フロー示すフローチャートである。

【0094】ブロックセクション処理によって、画像データが図形情報を持つ画像データであると、その画像データが、ハーフトーン、線画、折れ線等の図形タイプ

29

の認識を試みる。そのため、それらの図形ブロックは上述のような認識ができないので、図11に示される特定処理は、それらの図形ブロックを「未知図形」として表現する。

【0095】ステップS1101で、特定処理を実行するための図形タイプのブロックが選択される。ステップS1102で、そのブロックの図形タイプの補助属性が格納されている場合、選択された図形タイプのブロックの特定処理が実行される(ステップS1102)。このように、「ハフトーン」、「線画」、「折れ線」のいずれかの補助属性が格納されている場合(即ち、図10に示されるような処理によって補助情報が格納されている)、その図形タイプのブロックの特定処理は完了する。

【0096】一方、画像データが、ハフトーン、線画、折れ線のいずれかとして認識されない場合、そのブロックの図形タイプは未知として認識され、「未知」である補助属性をそのブロックに対応するノードに格納する(ステップS1104)。ステップS1104で、図形情報を持つ画像データの特定処理を中断する。図13は、処理対象のブロックが複数の画像データのタイプを含む合成領域を持つ画像データとして認識された場合(ステップS634)に、実行される特定処理の処理フローを示すフローチャートである。

【0097】つまり、ブロックセレクション処理によって、そのブロックは2つ以上の画像データのタイプを持つ画像データのブロックとして認識される。本発明の実施の形態によれば、複数の画像データのタイプを持つ画像データの認識を行うために合成領域がブロックセレクション処理の処理過程に適用される。ステップS1301で、ブロックの中に同一タイプの画像データを持つブロックが存在するか否かが判定される。例えば、図14Aに示されるように、原稿ページ1300には、タイトル1301、テキスト領域1302、1303、図形領域1305、1307、1309を含む。また、テキストキャプション1306も含んでいる。このような図14Aに示されるような原稿ページにおいて、ブロックセレクション処理を実行すると、図15Aに示されるような階層ツリー構造を持つ、分割された各領域のブロックが生成される。

【0098】また、図14Bに示すように、タイトル1300とテキスト領域1302、1303、可視ブロック1319、1320、1321として扱われる。しかしながら、テキストキャプション1306を持つ図形領域1305、1307、1309は、2つ以上の画像データのタイプを持つ画像データである「合成領域」1323としてグループ化される。

【0099】ステップS1301で、同一ブロックが存在しないか認識された場合、ステップS1302に進む。しかしながら、同一ブロックが認識された場合、ス

30

テップS1303に進み、そのブロックは「合成領域」として認識され、合成領域ノードが階層ツリー構造に付加される(ステップS1304)。ステップS1305で、それぞれの同一ブロックに対し、それらのブロックに対応する階層ツリー構造のノードの位置が変更され、「合成領域」ノードの下に付加される。図15Bは、上述の変更によって、合成領域ノード1323より分岐して生成されるテキスト1306、図形1305、1307、1309のノードを示す図である。

【0100】再び、図6Cの説明に戻る。原稿ページの階層ツリー構造が生成されると、ステップS640で、ステップS605で格納された確保された階層ツリー構造の記憶領域にその生成された階層ツリー構造が格納される。尚、各ブロックを表現するための最も簡単な方法は長方形であり、時にはブロックは孤立する長方形領域に存在しない。また、長方形のブロックの重なりが生じたりする。このブロックの重なりは、画像データのブロックの抽出の際にユーザを混乱させてしまう。そのため、長方形の使用するために、各ブロックを表現するために他の表現方法を提供する。他の表現方法は、連結成分の輪郭検出によって行われる。但し、輪郭のデータは簡潔化された連結する輪郭の組あるいは後述の連結する重ならない各ブロックの長方形であるとする。このように作られた輪郭は、ブロックの中に2つ以上の連結成分が存在する場合のそのすべてを含む連結成分から生成される。この方法は、連結する輪郭の組あるいは連結する重ならない長方形によって簡潔に、各ブロックを表現することができる。このような簡潔にブロックを表現する方法に基づいて、各ブロックの画像データは、各輪郭の組あるいは長方形の領域より直接抽出することができる。

【0101】ステップS641で、ブロックセレクション処理プログラムはユーザに対し次の原稿ページの解析を行うか否かを要求する。入力された原稿ページがない場合、ステップS642で、ブロックセレクション処理を中断する。一方、次の原稿ページの解析の指示がユーザからある場合、ステップS643で、階層ツリー構造の格納するために割り当てられた記憶領域に、充分な次の原稿ページの階層ツリー構造を格納する領域があるか否かを判定する。

【0102】階層ツリー構造の格納するために割り当てられた記憶領域に、充分な次の原稿ページの階層ツリー構造を格納する記憶領域がある場合、ステップS603に戻って、次の原稿ページのブロックセレクション処理を実行する。一方、階層ツリー構造の格納のために割り当てられた記憶領域に、充分な次の原稿ページの階層ツリー構造を格納する記憶領域がない場合、ステップS646で、エラーコードがユーザに対して出力される。

【0103】以上説明したように、本実施の形態によれば、ブロックセレクション処理プログラムは、入力した

31

原稿の有無に応じて処理を自動的に中断したり、階層ツリー構造の格納のために割り当てられた記憶領域に、充分な次の原稿ページの階層ツリー構造を格納する記憶領域がある場合、次の原稿ページの階層ツリー構造を格納する記憶領域を割り当てることができる。

【0104】尚、光学的文字認識(OCR)は、上述のブロックセレクション処理による出力を得るための、テキストブロックを読み取るために必要なアプリケーションとして限定されない。原稿の画像データを得ることのできる他のアプリケーションを用いて、例えば、原稿の格納と伝送、ブロックの認識と実行、原稿認識、原稿検索等のブロックセレクション処理を実行しても良い。

【0105】例えば、ブロックセレクション処理を使用する際に、より効率的に原稿の格納と伝送を実現するために、複数の画像データの圧縮方法を用いることでそれが実現できる。このように、例えば、テキストブロックは2値圧縮によって格納し、図形ブロックは、複数のレベルの図形圧縮によって格納することができる。また、ブロックの認識と実行は、ある形と他の形を区別するために各形に特定点を配置する必要をなくすためにブロックセレクション処理が生成される。ブロックセレクション処理を用いることで、ブロックの認識と実行は、上述のようなより一般的な方法で実行され、ブロックセレクション処理後、階層ツリー構造によって表現される。その階層ツリー構造は、新たな階層ツリー構造が存在する毎に、その内容が更新される階層ツリー構造の辞書を比較することによって認識される。

【0106】また、原稿の認識は論理的な解析によって実行される。基本的には、原稿認識は、所定の規則に従って原稿の1ページ分の画像データに分類する。そのため、ブロックセレクション処理の階層的な特徴により、ページ単位の物理的な解析を手伝うために原稿の構造的な基礎が獲得される。また、原稿の検索と格納は、上述のブロックセレクション処理に最大限使用することができる。ブロックセレクション処理を用いることで、原稿1ページ分の画像データは、上述の図3に示されるような形式で、その原稿画像よりもむしろ複数のブロックで表示される。その結果、ユーザは原稿に対するブロックの形式を一瞥することで、原稿の構成で必要とされるか否かについての視覚的な見当を与えるレイアウト情報を獲得することができる。そのため、ユーザは予め設定されている原稿の形式やある特定の要求に基づいて、テキストあるいは非テキストブロックの情報を獲得することができる。例えば、図16に示されるような標準ビジネスメモは、[To] 領域、[From] 領域、[CC] (カーボンコピー) 領域のようなメモから構成されている。図3の原稿において、各ブロックに対して特定の処理を指示する方法として、このビジネスメモを用いて、ユーザは、「3番目の線」に「CC」を指示するというような、各ブロックの位置情報に基づいて要求の仕方が可能

32

である。尚、本実施の形態では、ブロックの位置の情報によって処理対象のブロックを特定しているが、これに限定されず様々なブロックを特定できる情報によってブロックを特定しても良い。例えば、ユーザは、「3番目の線」と指示するかわりに、「原稿を半分にした時の、その原稿の上部の最終線よりも上に位置する水平線」というような特定方法で、対象とするブロックに「CC」を指示することができる。このように、ユーザによるブロックの特定は、自動的にブロックを抽出できる領域として解釈される。

【0107】尚、本発明は好適な施の形態について説明されたが、本発明は開示された実施の形態に限定されるものではないことが理解されるべきである。逆に、発明の精神と範囲内に含まれる、様々な修正や同等の変形が可能であるよう意図されている。特許請求に範囲の記載は、そうして修正及び同等の構成及び機能の全てに渡るように解釈が与えられるものである。

【0108】尚、本発明は、複数の機器から構成されるシステムに適用しても、1つの機器からなる装置に適用しても良い。また、本発明はシステム或は装置にプログラムを供給することによって実施される場合にも適用できることは言うまでもない。この場合、本発明に係るプログラムを格納した記憶媒体が、本発明を構成することになる。そして、該記憶媒体からそのプログラムをシステム或は装置に読み出すことによって、そのシステム或は装置が、予め定められた仕方で動作する。

【0109】

【発明の効果】以上の説明からも明らかなように、本発明によれば、ブロックセレクション処理の処理対象である原稿のデータ量と歪みを監視することで、処理精度が向上するブロックセレクション処理を実行できるページ解析システムを提供できる。

【図面の簡単な説明】

【図1】原稿ページの一例を示す図である。

【図2】図1の原稿ページに関して、ブロックセレクション処理の処理結果より得られる階層ツリー構造の一例を示す図である。

【図3】図2の階層ツリー構造に基づいて生成される原稿ページの表示例を示す図である。

【図4】本発明を適用する実施の形態の装置の外観図である。

【図5】図4の装置の構成を示すブロック図である。

【図6A】本発明のブロックセレクション処理において生成されるページ解析システムの処理フローを示すフローチャートである。

【図6B】本発明のブロックセレクション処理において生成されるページ解析システムの処理フローを示すフローチャートである。

【図6C】本発明のブロックセレクション処理において生成されるページ解析システムの処理フローを示すフロ

一チャートである。

【図 6 D】本発明のブロックセレクション処理において生成されるページ解析システムの処理フローを示すフローチャートである。

【図 7 A】入力原稿ページも一部分の領域の解析方法の処理フローを示すフローチャートである。

【図 7 B】解析された原稿の各領域の一例を示す図である。

【図 8】解析対象のピクセル画像データを格納する記憶領域の割当方法を説明するための図である。

【図 9 A】連結成分の 8 方向検索を示す図である。

【図 9 B】白輪郭の 4 方向検索を示す図である。

【図 10】非テキスト連結成分の区別処理フローを示すフローチャートである。

【図 11】図形ブロックの図形タイプの判定方法の処理フローを示すフローチャートである。

【図 12 A】折れ線を含む原稿ページの一例を示す図である。

【図 12 B】折れ線を含む原稿ページの一例を示す図である。

【図 12 C】折れ線の一例を示す図である。

【図 12 D】折れ線の一例を示す図である。

【図 13】合成領域の処理方法の処理フローを示す図で

ある。

【図 14 A】合成領域を含む原稿ページの一例を示す図である。

【図 14 B】合成領域の含む原稿ページの一例を示す図である。

【図 15 A】図 14 A の原稿ページの階層ツリー構造を示す図である。

【図 15 B】図 14 B の原稿ページの階層ツリー構造を示す図である。

10 【図 16】原稿からテキストを抽出するための標準的なメニューの一例を示す図である。

【符号の説明】

5 2 0 CPU

5 2 1 コンピュータバス

5 2 2 スキャナインタフェース

5 2 3 プリントインタフェース

5 2 4 ネットワークインタフェース

5 2 6 FAX/モデムインタフェース

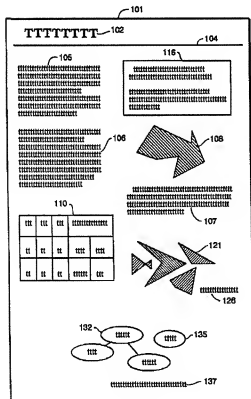
5 2 7 ディスプレイインタフェース

20 5 2 8 キーボードインタフェース

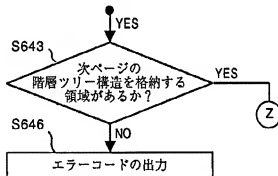
5 2 9 マウスインタフェース

5 3 0 メインメモリ

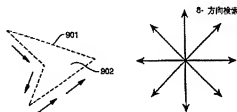
【図 1】



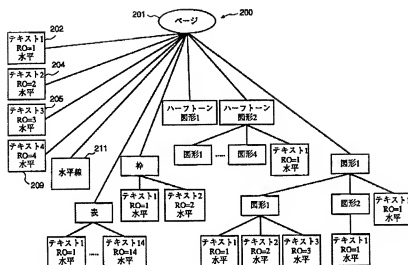
【図 6 D】



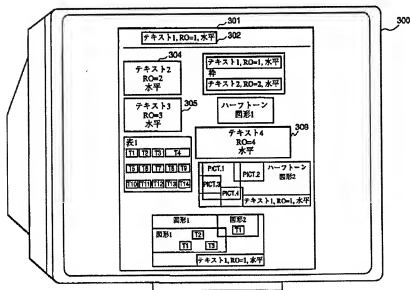
【図 9 A】



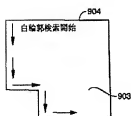
【図2】



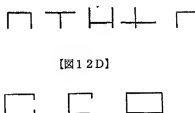
【圖3】



【圖 9 B】

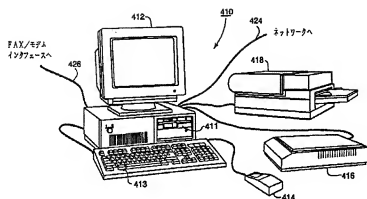


【图 12C】

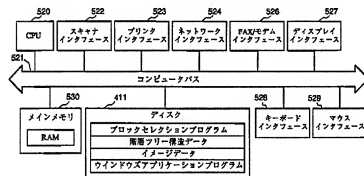


【图 1 2 D】

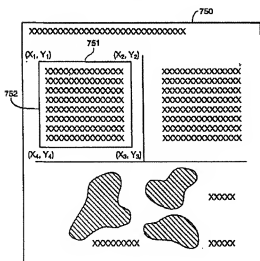
【図 4】



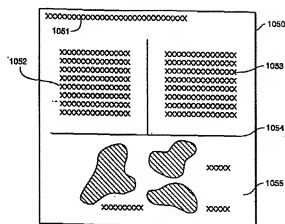
【図 5】



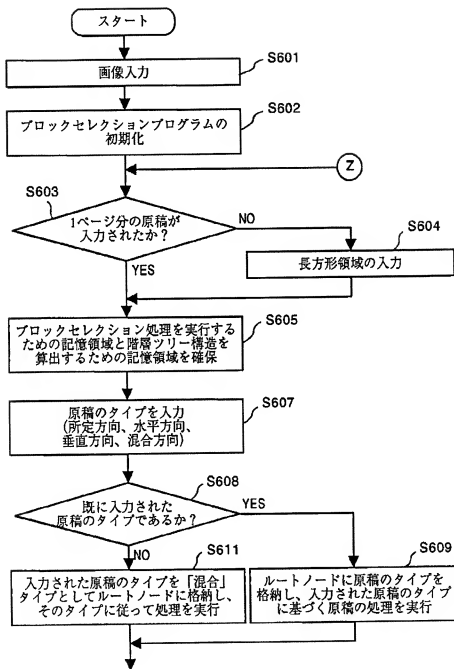
【図 7 B】



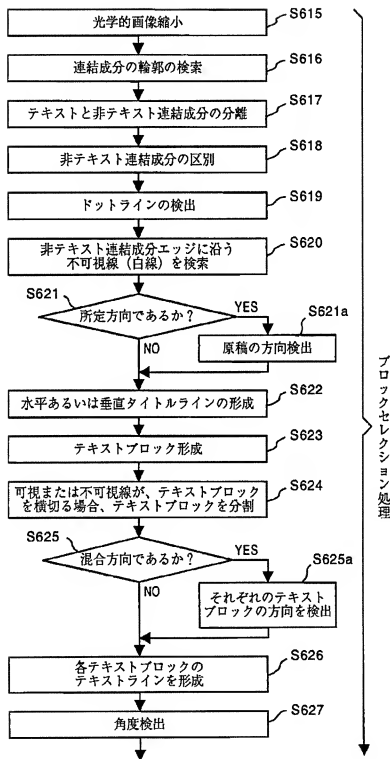
【図 12 A】



【図 6 A】

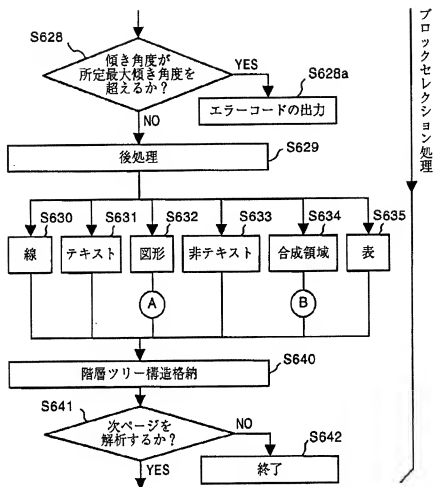


【図 6 B】

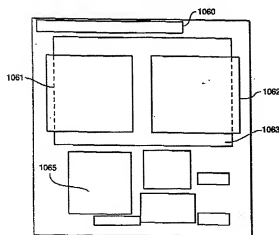




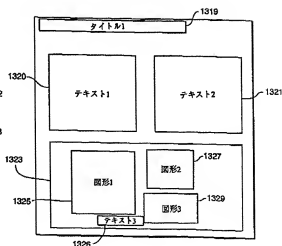
【図6C】



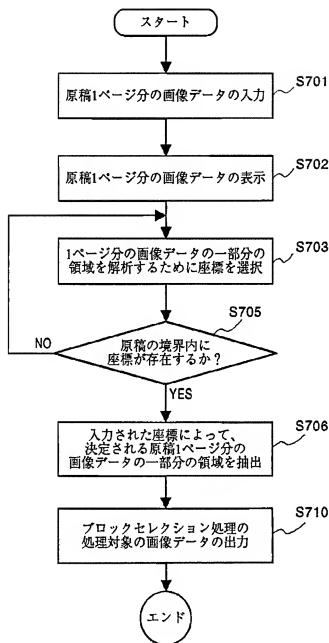
【図12B】



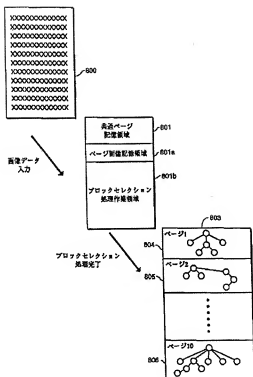
【図14B】



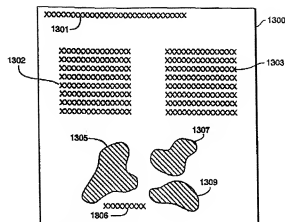
【図 7 A】



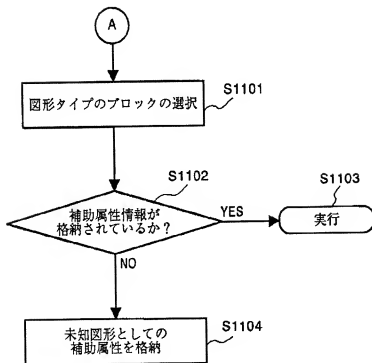
【図 8】



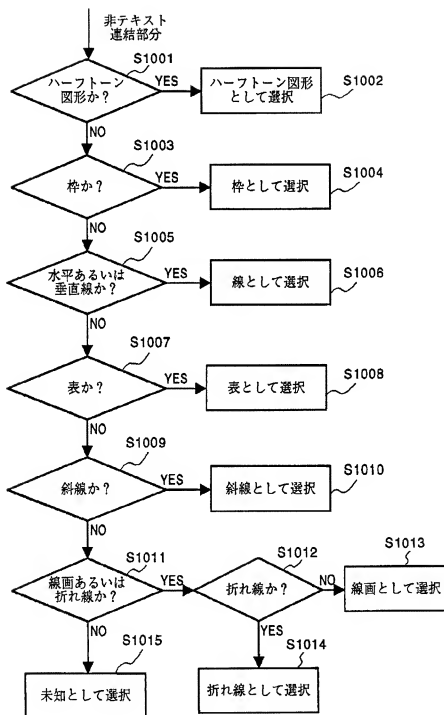
【図 14 A】



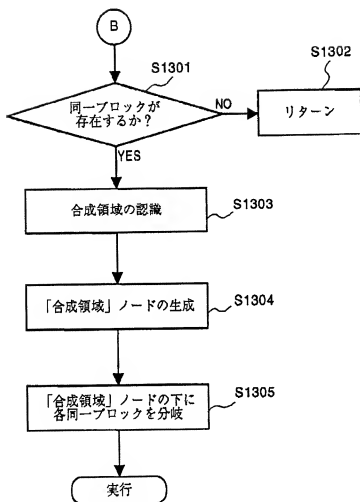
【図 11】



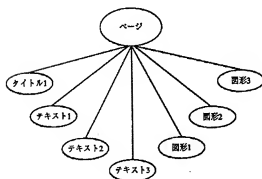
【図10】



【図13】



【図15A】



【図15B】

